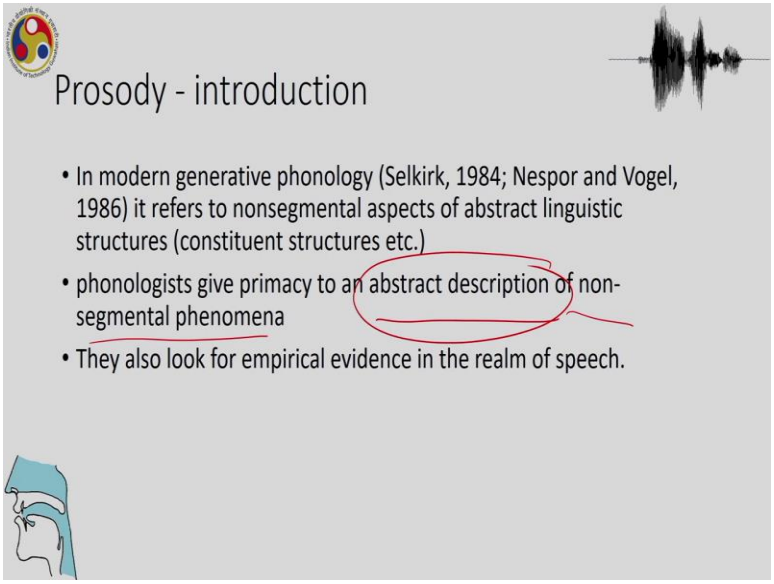
Phonetics and Phonology: A Broad Overview
Professor Shakuntala Mahanta
Department of Humanities and Social Science
Indian Institute of Technology Guwahati
Lecture 25
Phonetics and Phonology of Intonation, Microprosody, Stylization

Welcome to this NPTEL MOOC course on phonetics and phonology a broad overview. So, we are in the unit, that is, the last unit of tone and intonation. Today we will look at what do we mean by prosody and its linkage with intonation.

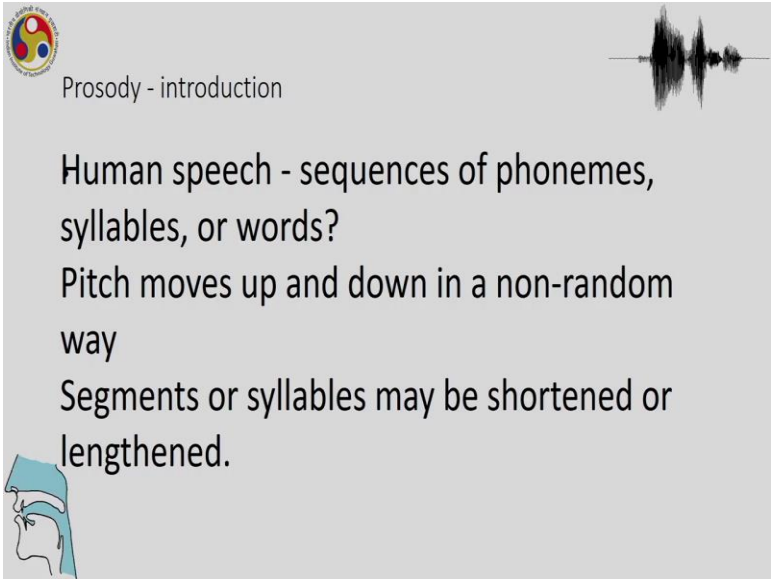
(Refer Slide Time: 00:52)



So, we will start with prosody and what it means. In modern generative phonology, the reference is given here, which we will put at the end of the course, refers to non-segmental aspects of abstract linguistic structures. So, phonologists give primacy to abstract descriptions, as we already know from our introduction to phonology, and in prosody also phonologists try to look at the abstract representations and the abstract descriptions of even non-segmental phenomena.

So, as you understand, in prosody we talk about non-segmental phenomena. And they also look for empirical evidence in the realm of speech. So, while they look for abstract descriptions or abstract generalisations or abstract analysis, they also look for empirical evidence in the realm of speech.

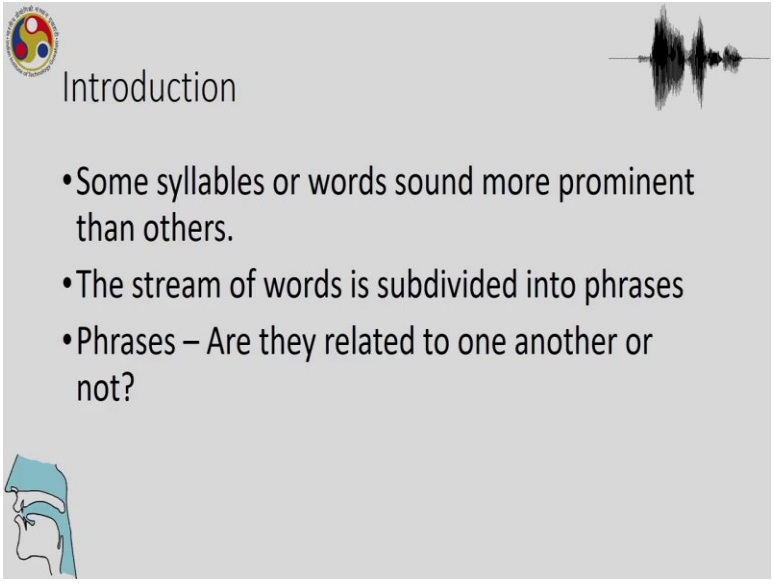
(Refer Slide Time: 01:52)



And while we talk about prosody, we need to know about human speech. What are the sequence of phonemes syllables or words? So, what happens in the level of pitch, pitch moves up and down in a non-random way, so it will be phonologists and linguists for a long time known that when we make use of pitch in speech, we do not use it randomly.

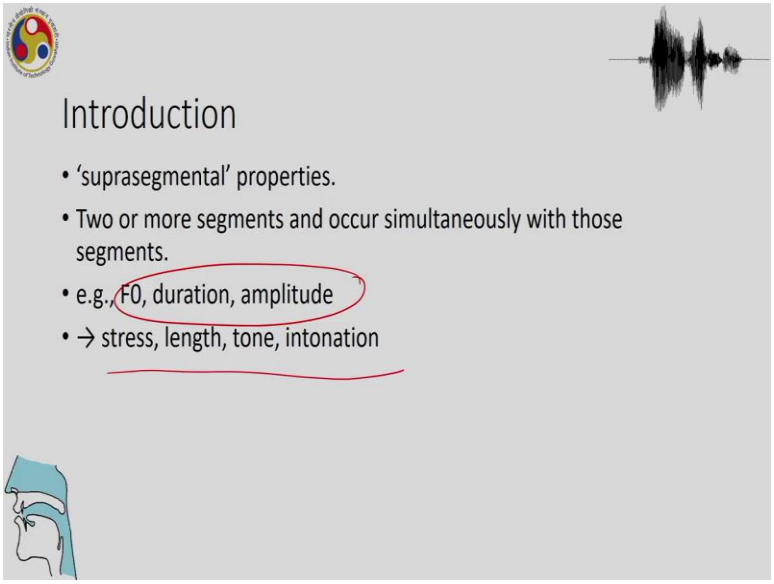
So, there is a reason why when it rises or it falls or there is a break or there is a pause, or something is lengthened. All those things that we normally will not study in the segmental part of phonology and phonetics, we will study in prosody. So, segments or syllables may be shortened or lengthened.

(Refer Slide Time: 02:48)



And some syllables or words sound more prominent than others. And the stream of words is subdivided into phrases. So, we will come to this, we will talk about this more when we look at the phonology of intonation. And what are phrases, do we actually speak in such a way that we break the utterances into parts.

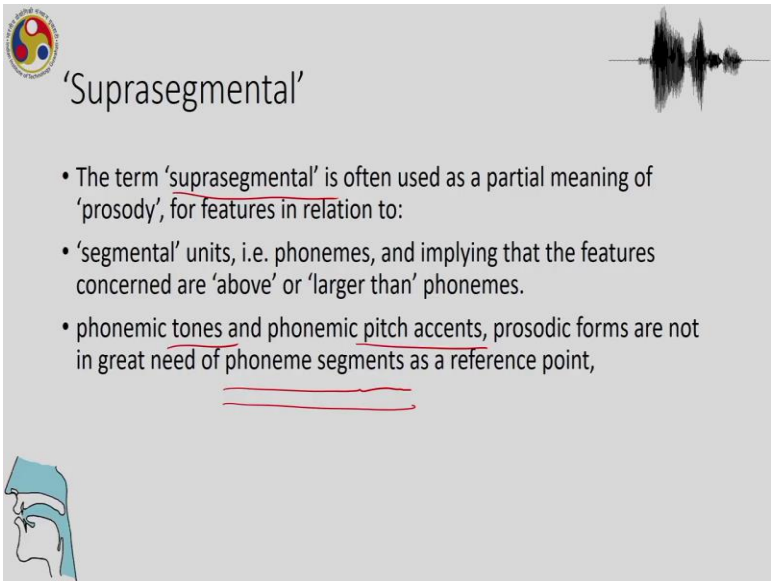
(Refer Slide Time: 03:16)



So, very often, the word suprasegmental is used to describe prosody, when two or more segments occur simultaneously with those segments. So, these are the properties which are often thought to be suprasegmental, that is, above the level of the segment, that is fundamental frequency,

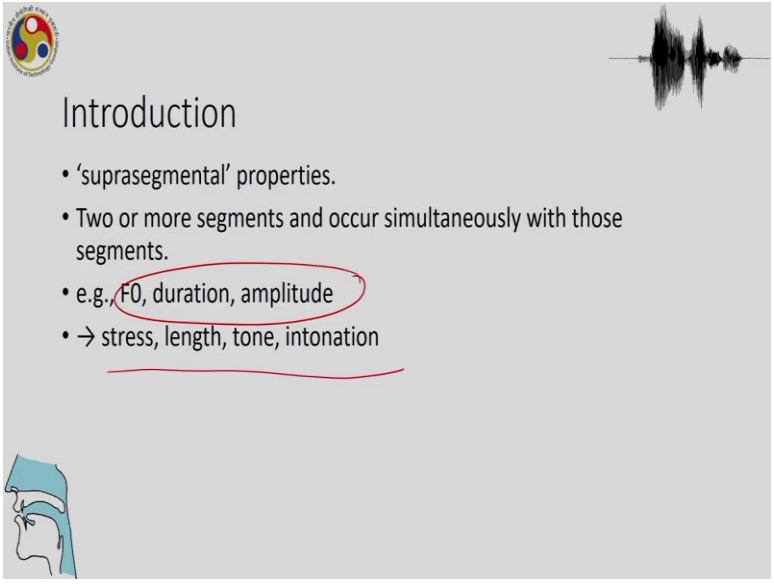
duration, amplitude and also stress which we studied, which we looked at in the previous lecture, stress and length, length of a vowel, word length and various aspects related to length in terms of its relation, in terms of its position in a word and also tone and intonation in general.

(Refer Slide Time: 04:03)



So, the term suprasegmental is often used as a partial meaning of prosody for features in relation to segmental units, that is, phonemes, and implying that the features concerned above or larger than phonemes.

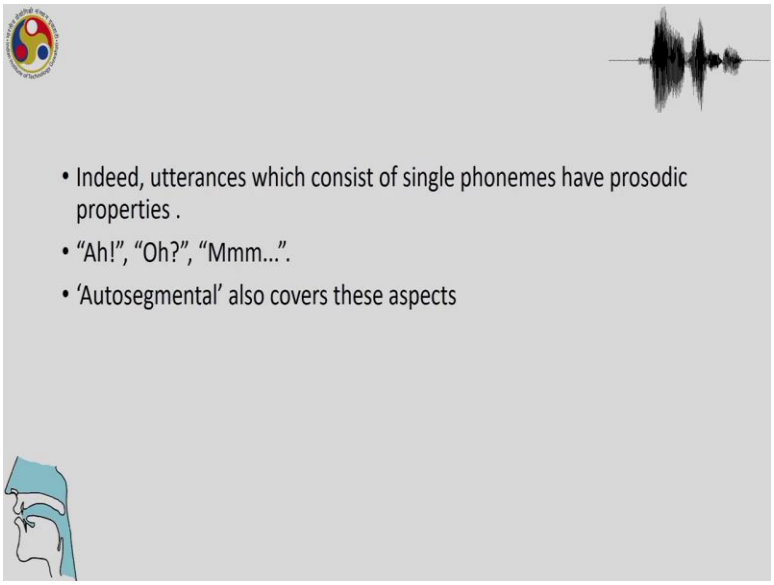
(Refer Slide Time: 04:19)



And what are those features these are the features that is fundamental frequency, duration, amplitude, that these are above, all these constitute larger units than the segments that have studied so far.

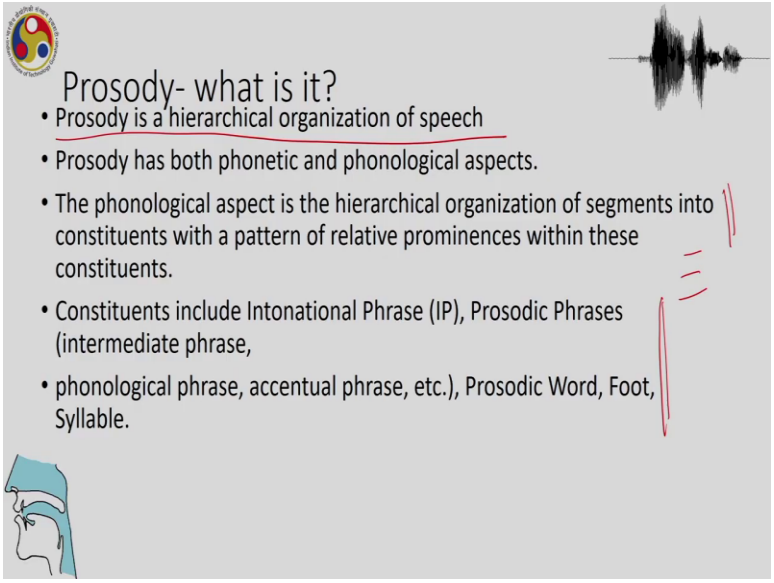
And phonemic tones and phonemic pitch accents and prosody forms are not in great need of phoneme segments as a reference point. So, phoneme segments, they do not just look for one phoneme segment that is tones and pitch accents, they are not actually dependent on particular phoneme segments as a reference points and that they are above and that is why the name suprasegmental. So, it is sometimes used as one of the meanings of prosody.

(Refer Slide Time: 05:17)



But as we know that utterances with single phonemes also have prosodic properties. So, how do we know that? Because they are not related to more than one segment, they can be just related can be expressed with just one segment like Mmm, Ah. So, those are just one segments, and those express something in the prosody, so there will be surprise or question et cetera, they do not need more than one segment.

(Refer Slide Time: 05:50)

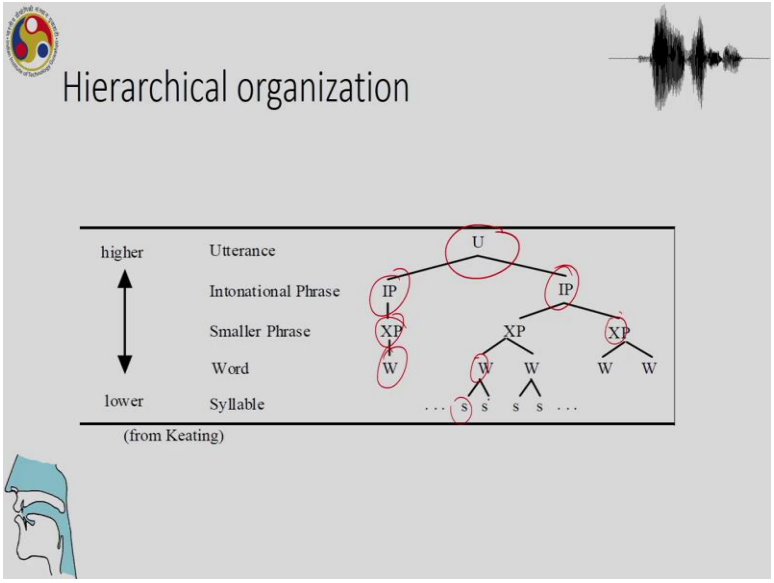


So, in a sense, then prosody captures this, the idea of prosody captures this that it is not just more than one segment, that this is we are talking about the level of organisation, which is not really dependent on the segments. So, prosody is a hierarchical organisation of speech, when we talk about phonology, it is inherent in the understanding in phonology is that prosody is a hierarchical organisational speech and prosody has both phonetic and phonological aspects.

The phonological aspect is hierarchical organisation of segments into constituents with a pattern of relative prominences within these constituents. So, phonological aspect is the hierarchical, this aspect, that is the hierarchical aspect that there are levels of organisation of prosody and that there is relative prominence. So, which means one is more prominent than the other in a sequence. And those aspects are extremely important when you are looking at the phonology of prosody.

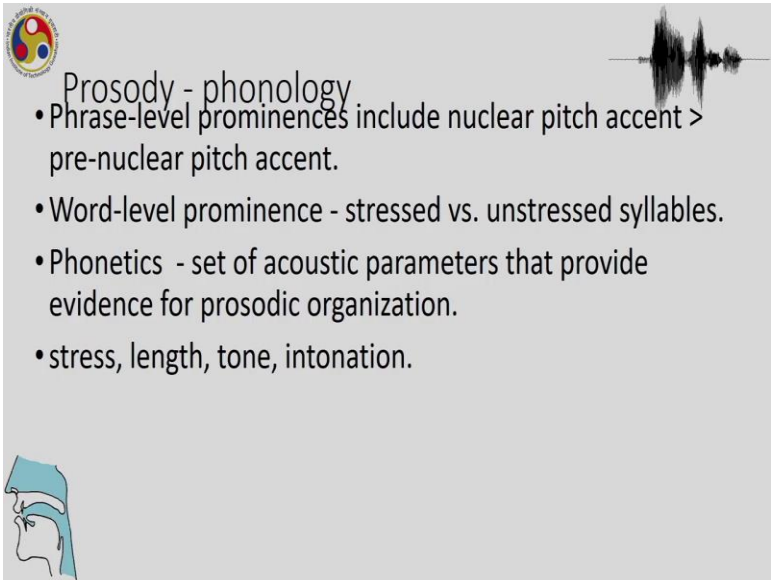
And further, that the chunking aspect that we talked about before that constituents include international phrases, prosodic phrases, phonological phase, accentual phrase, et cetera. We will talk about these in greater detail in the other lecture on the international phonology.

(Refer Slide Time: 07:19)



So this, as you see here, we can see the hierarchical organisation, that this is the largest level which is the utterance that is the sentence. And then we organise it into further units like the international phrases and they can be two international phrases in utterance they can be smaller phrases and then there could be the word level and then the word can be further divided into other parts like the syllables, like feed, et cetera.

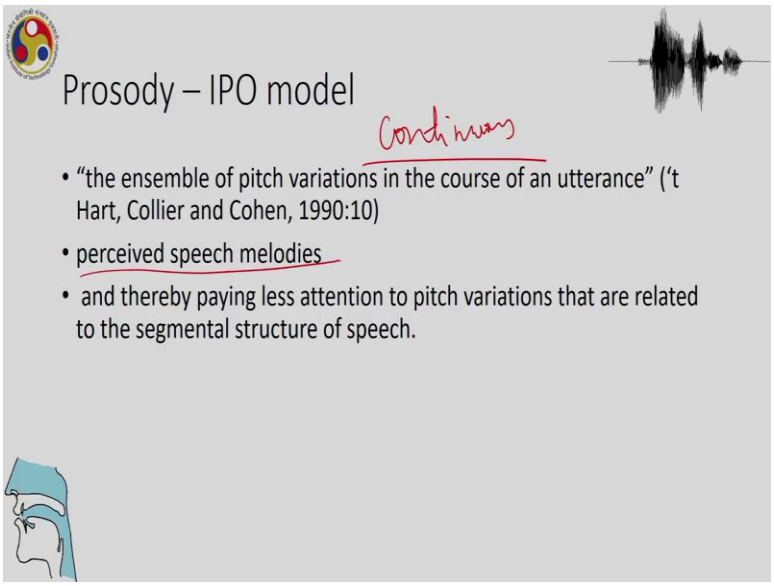
(Refer Slide Time: 07:57)



So, phrase level prominences include nuclear pitch accent and pre-nuclear pitch accents. We will talk about why some analysis have nuclear pitch accents and pre-nuclear pitch accents in the

lecture where we talk about international phonology. And however, in every analysis of intonation, word level prominence, that is stress versus unstressed syllables, is very important. And when we talk about the phonetics of prosody, set of acoustic parameters that provide evidence for prosodic organisation, that is stress, length all these things are important in the phonetics.

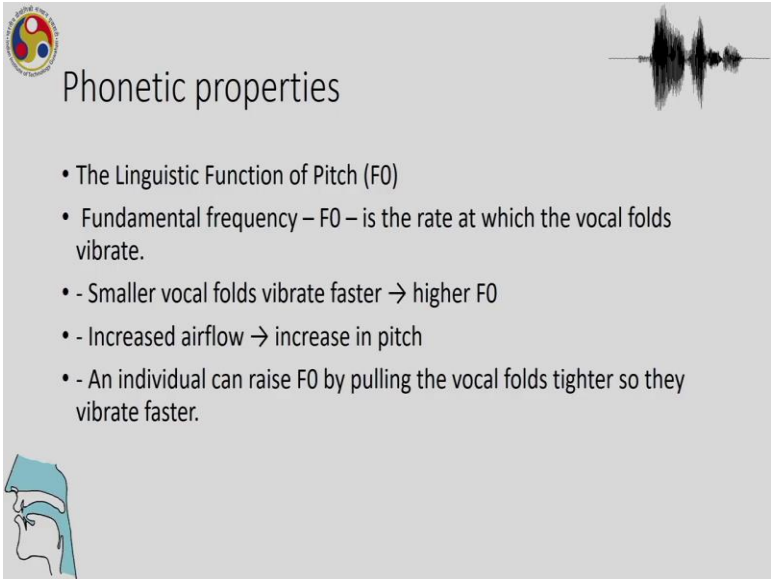
(Refer Slide Time: 08:43)



So, the IPO model is one such model which models prosody. So, we are going to talk about the IPO model in this lecture and we will talk about other models in the other lectures. So, in the IPO model, the ensemble of pitch variations in the course of an utterance. So, this is the definition of intonation. So, pitch variations in the course of utterance. Now this when you talk about the course of an utterance note that we are talking about the continuous changes in utterance. That is what you are talking about here.

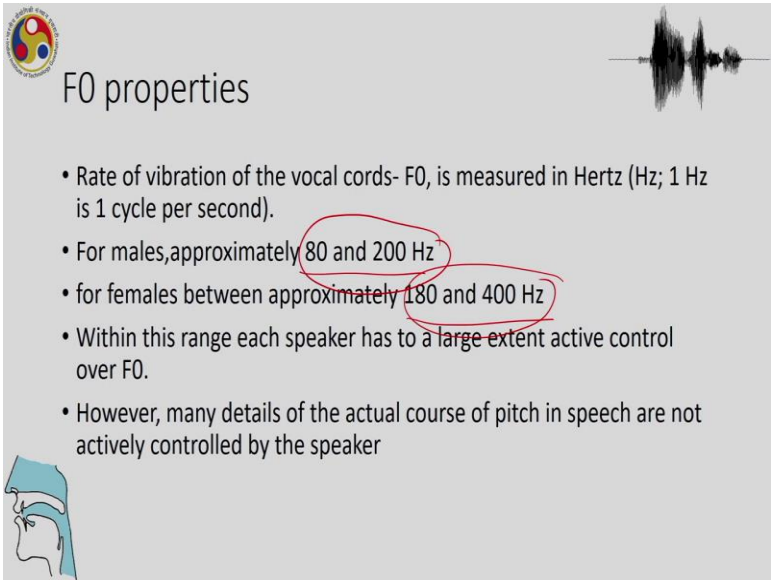
So, that is the difference between different approaches to prosody that some talk about the continuous changes, some talk about it in categorical way that these are different pitch accents and some are talking about the relative differences, continuous changes in the utterance. So, keep that in mind. And also the IPO model paid attention to the perceived speech melodies and thereby paying less attention to pitch variations that are related to the segmental structure of speech.

(Refer Slide Time: 10:08)



And the linguistic function of pitch, fundamental frequency is the rate at which the vocal folds vibrate. And we know this from our classes, from lectures on acoustic phonetics. So, fundamental frequency is the rate at which the vocal folds vibrate and smaller vocal folds vibrate faster and leads to increase airflow and increase in pitch. And an individual can raise fundamental frequencies by pulling the vocal folds tighter, so they vibrate faster.

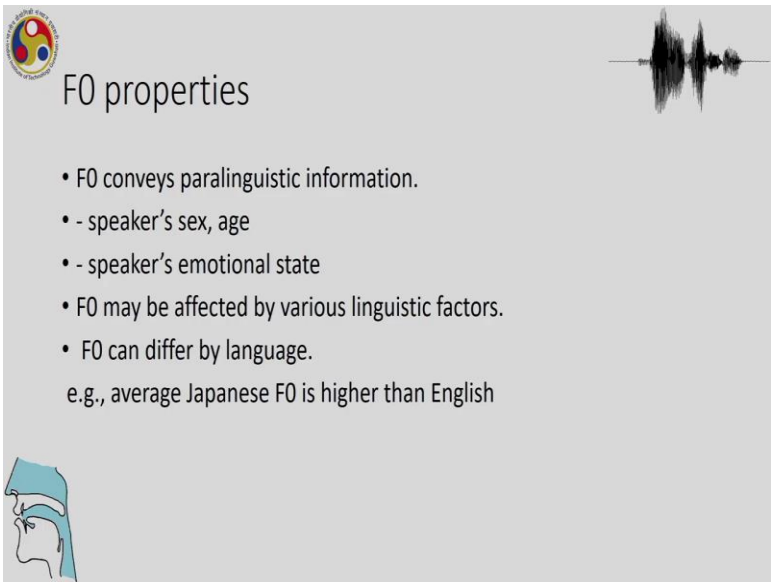
(Refer Slide Time: 10:39)



And the rate of vibration of the vocal folds, F0, is measured in hertz. And that is, as we know from the classes in acoustic phonetics, that 1 hertz is 1 cycle per second. And there is a

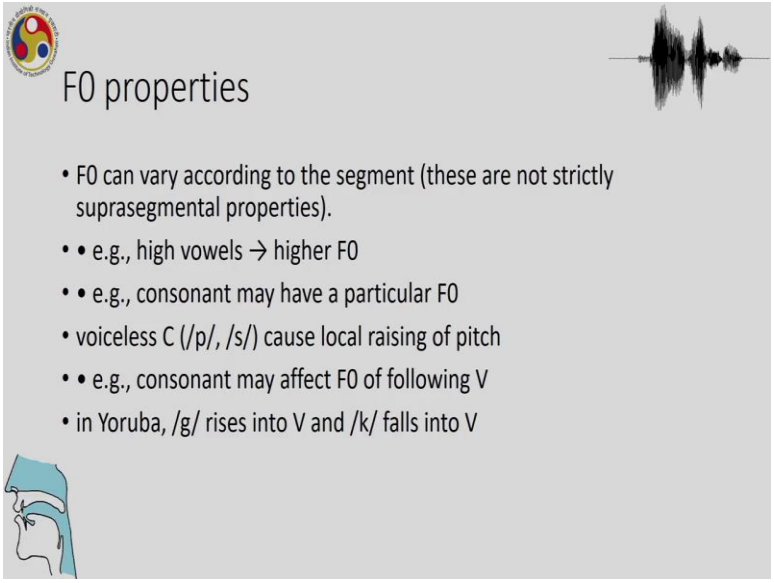
difference between generally, but this is not an absolute difference. So, generally for males it could be around 80 to 200 hertz and for females around 180 and 400 hertz, the pitch range along which we will find the pitch differences, the pitch changes. Within this range each speaker has large extent active control over the fundamental frequency. And however, many details of the actual cause of pitch in speech are not actively controlled by the speaker.

(Refer Slide Time: 11:33)



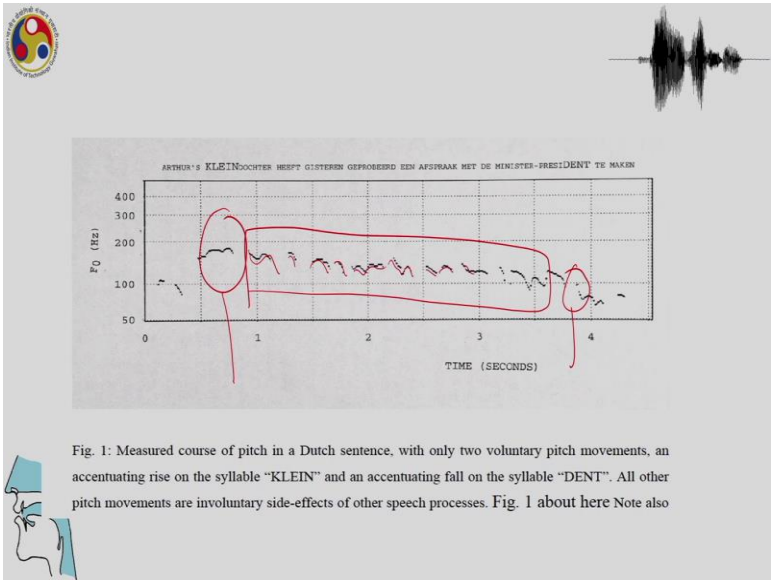
And fundamental frequency conveys paralinguistic information, our emotions, happiness, sadness, et cetera. The speaker's sex and age and speaker's emotional state, the fundamental frequency may be affected by various linguistic factors, fundamental frequency can differ by language. Average Japanese fundamental frequency is higher than that of English, et cetera. So, this fundamental frequency is dependent on a lot of factors. So, the variations will be dependent on those factors.

(Refer Slide Time: 12:14)



So, fundamental frequency can vary according to the segment also. And these are not strictly suprasegmental properties, high vowels have high fundamental frequency and consonants may have a particular fundamental frequency depending on whether they are voiceless or voiced. Voiceless consonants caused local pitch raising, consonants may affect fundamental frequency of following vowels. In Yoruba ga rises into a vowel, but ka falls into a vowel. So, this is either a rise or fall depends on whether they are voiced or voiceless.

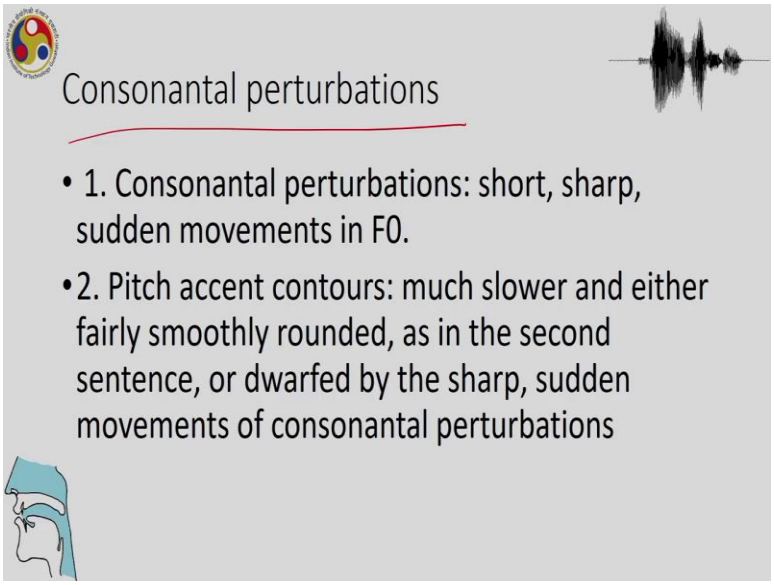
(Refer Slide Time: 12:54)



So, now, what you see here is a measured course of pitch in a Dutch sentence with only two voluntary pitch movements and accentuating rise in the syllable. So, note this. And then accentuating fall on dent. So, all the others are what we mean to say that all these changes here are involuntary changes. Only these two are controlled by the speaker.

And why is this so and what is it, is there a perceptual bearing of these kinds of changes is what we are going to talk about in the rest of the lecture. So, all these involuntary rise and falls that you see here in a pitch, that is falling, again rising and rising here falling. We are saying that all these rises and falls are involuntary.

(Refer Slide Time: 14:00)

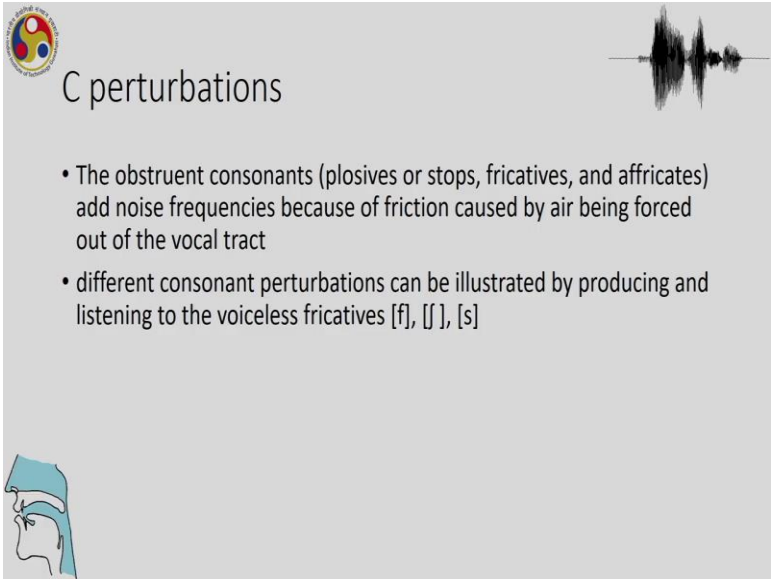


And these are called continental perturbations. So, continental perturbations are short, sharp, sudden movements in the fundamental frequency. And pitch accent contours a much slower and either fairly smoothly rounded, as in the second sentence, or dwarfed by the sharp, sudden movements of continental perturbations.

So what we are saying is that, as you can see, this is slow, smooth rise here, unlike this sharp sudden rise and fall. So, that is the difference between pitch accents or the ones the pitch change controlled by the speaker and the pitch changes which are not controlled by the speaker, this are the sharp sudden rise and falls.

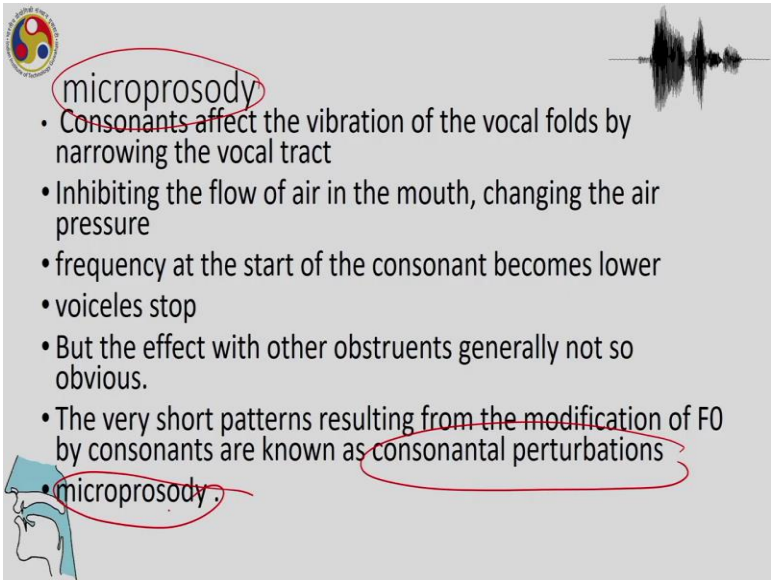
And the sharp sudden rise and falls are what we call continental perturbations.

(Refer Slide Time: 14:57)



And the obstruent consonants plosives or stops, fricatives and affricates add noise frequencies because of friction caused by air forced out of the vocal tract. We know fricatives are noisy sounds, we have discussed that in the acoustic phonetics lecture. And different consonantal perturbations can be illustrated by producing and listening to the voiceless fricatives f, sh, s. We have seen in the acoustic phonetics lecture how these cause different noises at different frequencies.

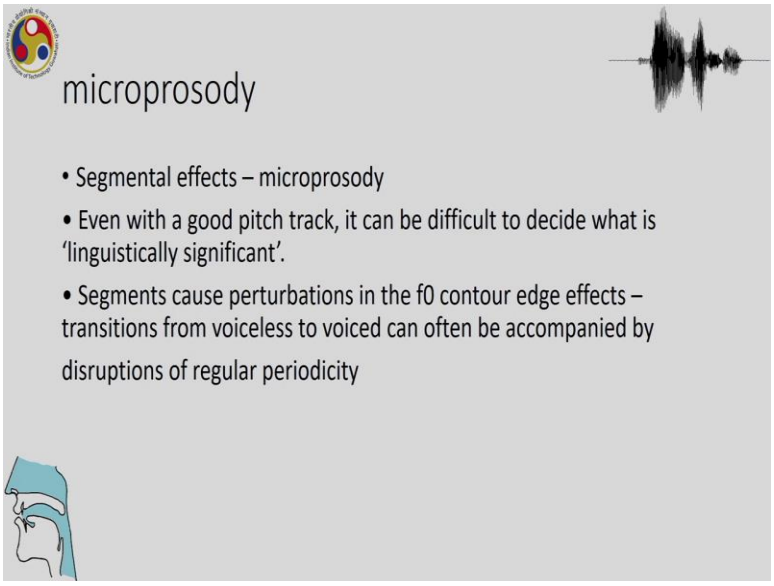
(Refer Slide Time: 15:30)



So, this is what we call the constant perturbations and other changes in the pitch contour is called microprosody. So, consonants affect the vibration of the vocal folds by narrowing the vocal tract, inhibiting the flow of air in the mouth, changing the air pressure and frequency the start of the consonant becomes lower for voiceless stops. But the effect with other obstruent's generally not so obvious.

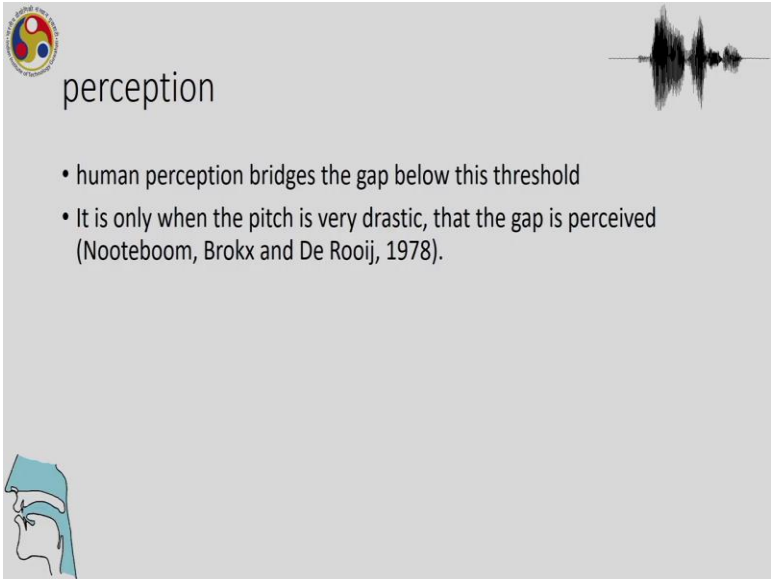
So, voiceless stops can affect the vowels, the pitch because it narrows the vocal tract inhibiting the flow of air, it changes the air pressure. But the obvious effect with other obstruent's generally not so obvious, as we just said. And the very short patterns resulting from the modification of fundamental frequency by consonants are known as consonantal perturbations or microprosody.

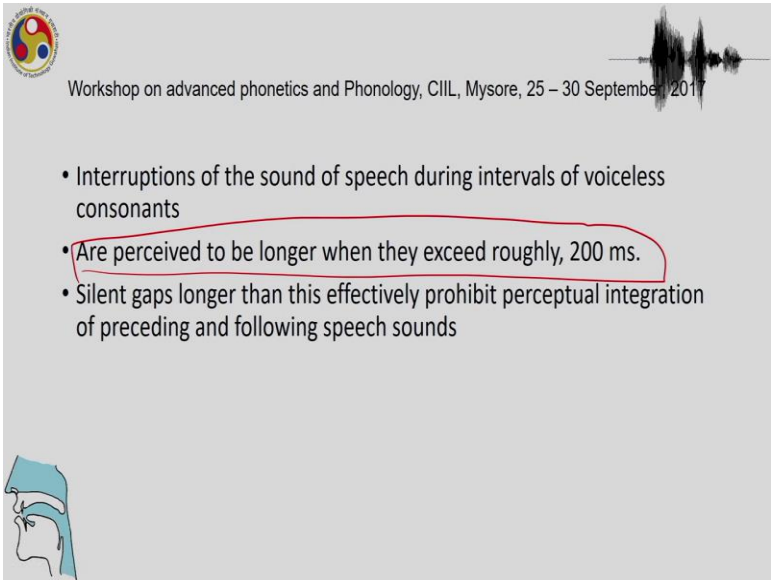
(Refer Slide Time: 16:35)



So, these segmental effects are called microprosody and even with a good pitch track it can be difficult to decide what is linguistically significant. Segments cause perturbations in the fundamental frequency, in the fundamental frequency come toward edge effects, transitions from voiceless to voice can often be accompanied by disruptions of regular periodicity.

(Refer Slide Time: 17:03)

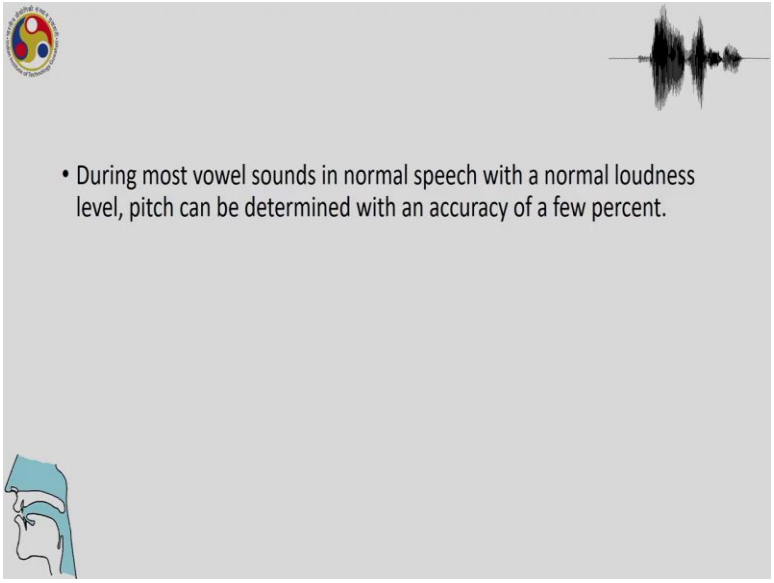




So, human perception bridges the gap between this threshold. So it is only when pitch is very drastic that the gap is perceived. But interruptions of the sound of speech during intervals of voiceless consonants are perceived to be longer when they exceed roughly 200 milliseconds. So, the human perception threshold compensates for these gaps and silent gaps longer than this effectively prohibit perceptual integration of preceding and following sound.

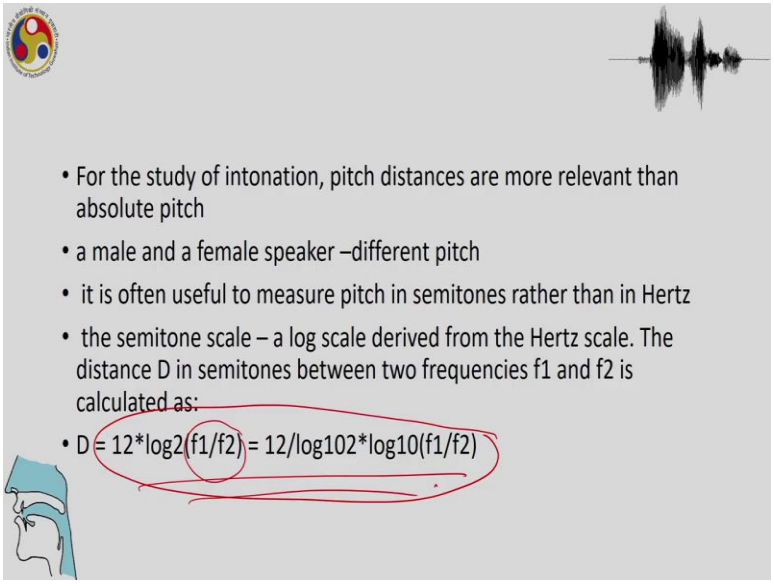
So, 200 milliseconds is a long time. So, something below that is not even perceived by the human perception mechanism. So, that is why these microprosodic effects are not relevant when we are talking about the perception of general prosody in human speech.

(Refer Slide Time: 18:08)



During most vocal vowel sounds, in normal speech with normal loudness, pitch can be determined with an accuracy of a few percent. So, normal loudness level, pitch is determined with accuracy but if there it is higher, if it is louder, then, that may interfere with perception.

(Refer Slide Time: 18:30)

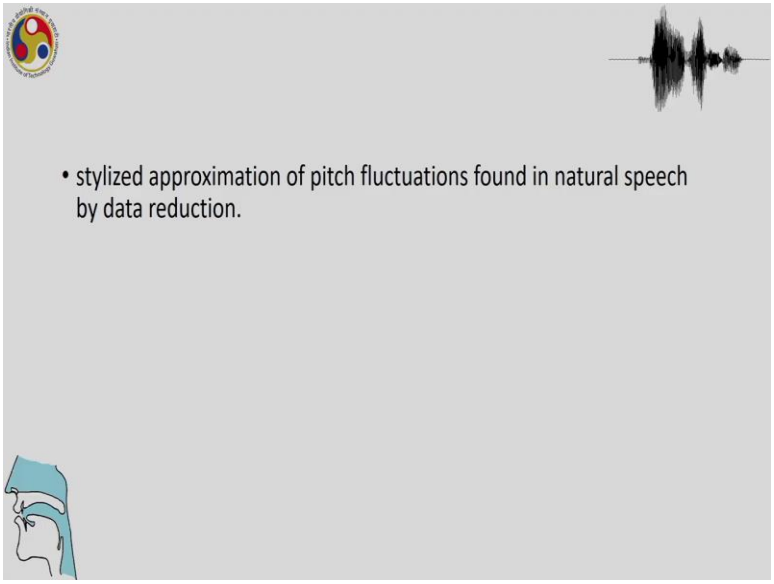


For the study of intonation, pitch distances are more relevant than absolute pitch. So, the distance from one height to another is more relevant. And for male and female speaker we have different pitch levels. It is often useful to measure pitch in semitones, because that gives you an idea about the perception. The semitones scale a log scale derived from the hertz scale, the distance D in

semitones between two frequencies f and 1 there is a formula for the distance and we can get the log value after dividing the f1 from the f2.

So, this is a formula for the calculation of semitone, the semitones scale. So we will have to find the logarithmic values after multiplying the f1 by f2. So, this is the calculation and we will not talk about semitone scale anymore, but if you are interested in perceptual studies of pitch, then the semitones scale can be useful.

(Refer Slide Time: 19:42)

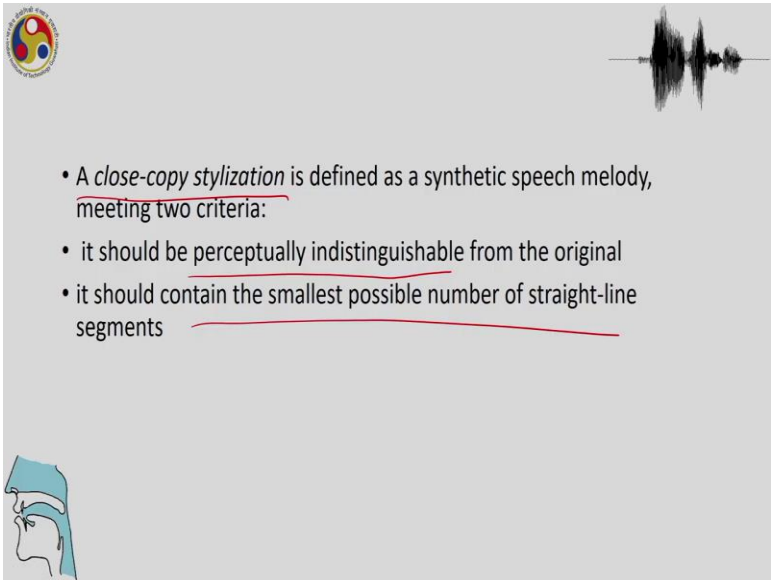


So, now we come to the stylized approximation pitch fluctuations in natural speech. So, why do we resort to or why would we have stylized approximation of pitch fluctuations? Because as we just talked about, that the actual, what you see in the fundamental frequency, the short sharp rises and falls are actually not even perceived by human speakers. So, that is why we have stylized approximation which takes away these small changes which are not perceived by speakers.

(Refer Slide Time: 20:28)

And that brings us to the perceptual model of pitch and this is called close copy stylization of pitch in speech utterances applied by De Pijper to British English. And so we see the natural measured F0 curve of an English utterance, we will see that close very soon.

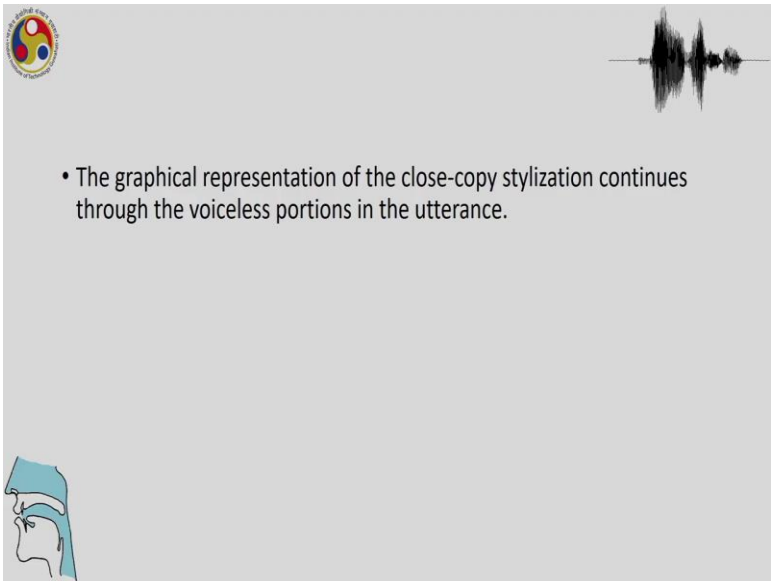
(Refer Slide Time: 20:46)



A closed copy stylization is defined as a synthetic speech melody meeting two criteria, it should be perceptually indistinguishable from the original, it should contain the smallest possible number of straight line segments. So, this is why it is called the perceptual module because it should be the close copy stylization or after removal of the micro prosodic effects it should be

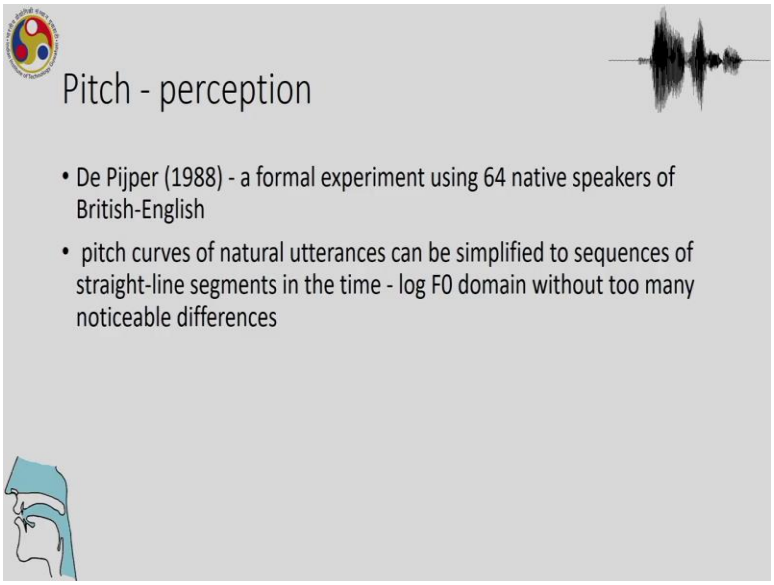
perceptually similar to what was the original with the micro prosodic effects. And it should contain the smallest number of straight line segments which means that if you have too many straight line segments, it will then also include the consonantal perturbations.

(Refer Slide Time: 21:38)



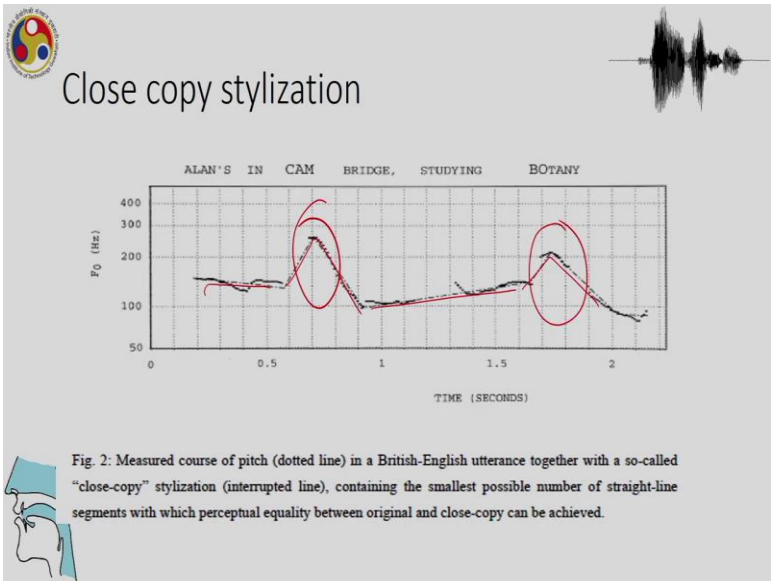
So, the graphical representation, the close copy stylization continues through the voiceless portions of the utterance. So, when we have voiceless portions, we do not have the F0 track at those points, but close copy stylization will ignore those and also draw a straight line through the voiceless portions.

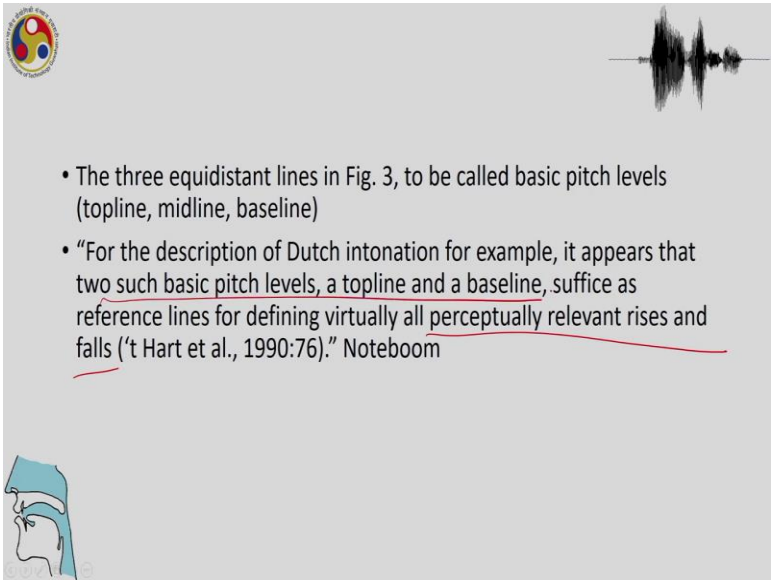
(Refer Slide Time: 21:58)



And De Pijper, a formal experiment using 64 native speakers of British English and the pitch curves of natural utterances can be simplified to sequences of straight line segments in a time log F0 domain without too many noticeable differences.

(Refer Slide Time: 22:17)





So, this is what we are talking about a while ago. This is the closed copy stylization, unlike what you had seen before with the consonantal perturbations. As you can see, this utterance has the smallest number of pitch rise and falls. You can basically just see two here. And what are those two, those are the parts which are emphasised by the speaker. And as we had in the beginning of the lecture, we had talked about how stress and relative prominence is important in

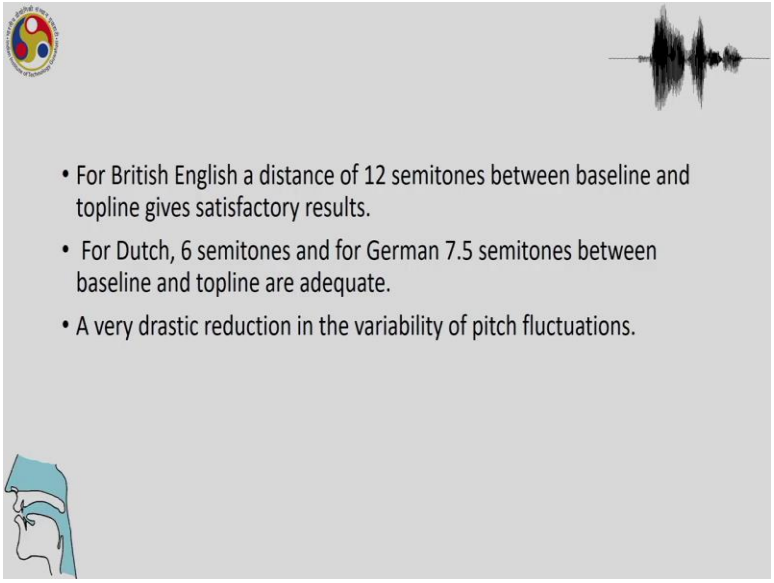
understanding pitch, you can see that here, that these are the two most important points in the utterance.

So, these measured cause of pitch in a British English utterance, together with so called closed copy stylization containing the smallest number of straight line segments with which perceptual equality between the original and close copy can be achieved. What is the sentence? The sentences are Alan's in Cambridge studying botany. So, here Cambridge and botany are the two syllables which are most prominent and the close copy stylization shows exactly that.

The three equidistant lines and figure 3 to be called basic patch levels, top line midline and baseline. And we can see that we had the baseline for the description and then what we have the prominence moves from the baseline and it shows that that those are the most sort of significant events in this utterance.

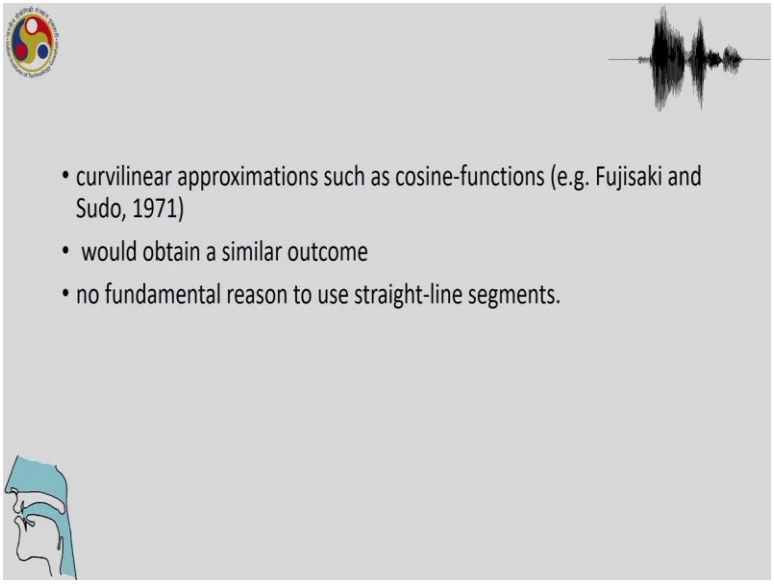
So, for the description of Dutch intonation, for example, it appears that two such basic pitch levels, top line and baseline suffice as reference lines for defining virtually all perceptually relevant rises and falls. So, to so top line and a baseline, so if you just take into account that this is the baseline, a little above 100 hertz, and that the top line moves all the way from 200, between 200 to 300 hertz and you have your most significant pitch events there, the prominent part.

(Refer Slide Time: 24:55)



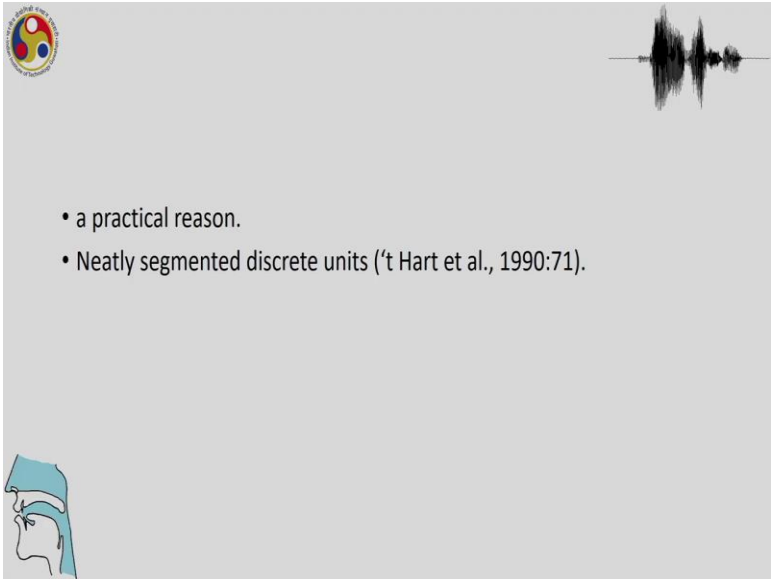
So, for British English a distance of 12 semitones between baseline top line gives satisfactory results for Dutch 6 semitones and for German 7.5 semitones, a very drastic reduction in the variability of pitch fluctuations. So, now, we can see that we need to understand that the pitch fluctuations that are happening which are controlled by the speakers are controlled because it is driven by certain aspects of prosody. Here we can see prominence and those are things which have the important things in the prosody, in the utterance which have to be captured by any prosodic model.

(Refer Slide Time: 25:43)



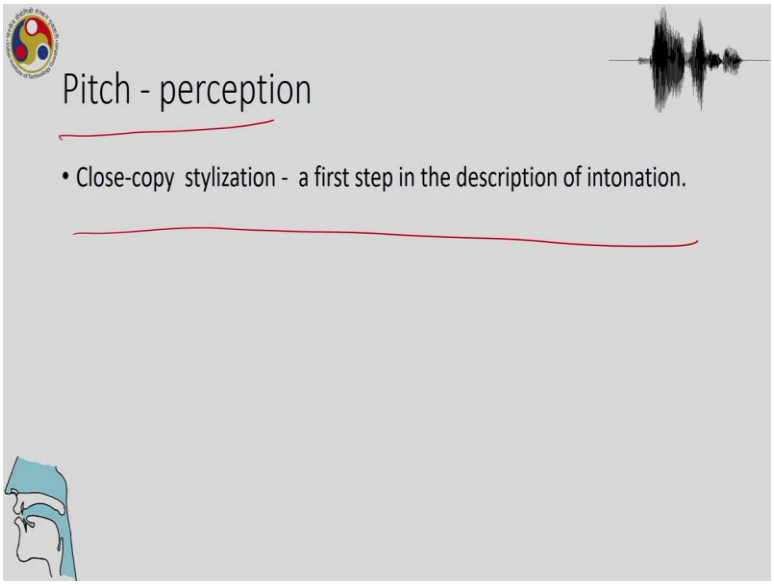
So, other models like the Fujisaki model with curvilinear approximate such as approximations which is cosine functions, will also obtain similar outcomes as in get the most important pitch movements. And no fundamental reason to use straight line segments, it is most important to capture the parts which are most prominent.

(Refer Slide Time: 26:10)



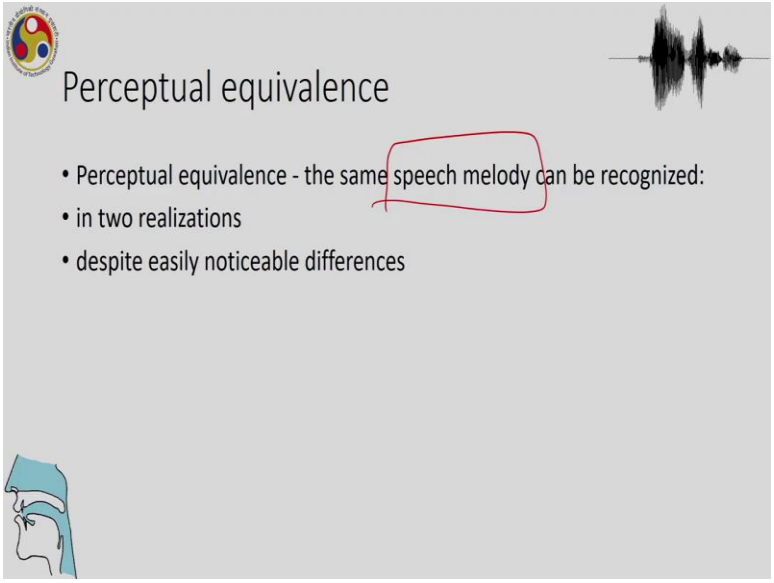
And then however, in the t hart model, we have neatly segmented discrete units.

(Refer Slide Time: 26:19)



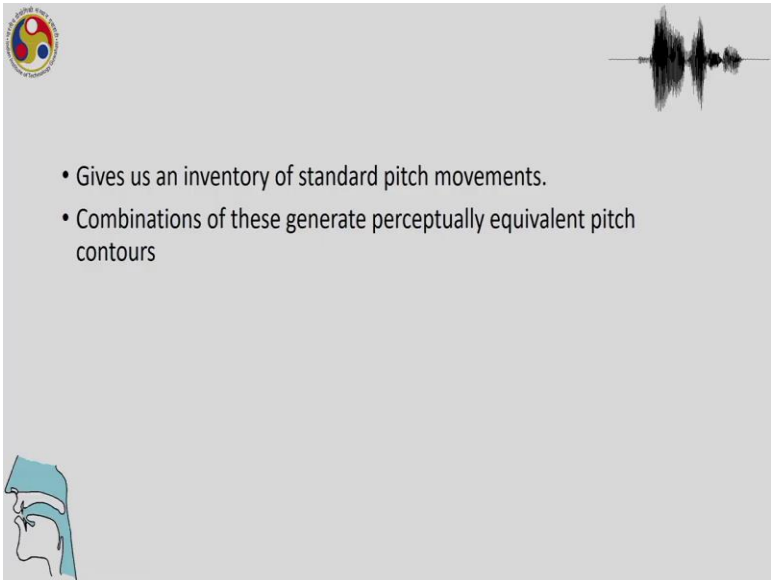
So, in an understanding perception of pitch, close study stylization was the first step in the description of intonation from the perceptual point of view.

(Refer Slide Time: 26:36)



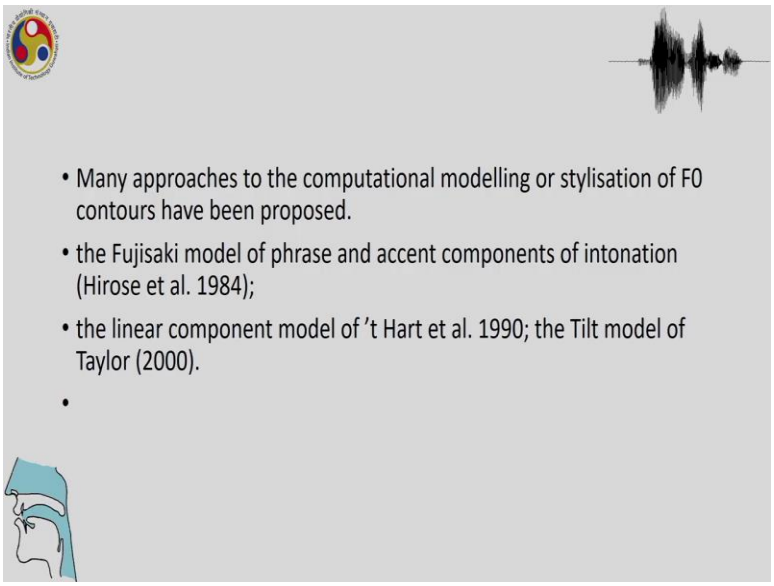
So, perceptual equivalence, the same speech melody can be recognised in two realisations despite easily noticeable differences. So, the same speech melody, so melody is something that is talked about in this type of an analysis, which talks about the continuous changes. We will talk about the categorical aspects of speech prosody in the phonology models. But here we can see that the pitch changes have been captured, the perceptual relevance and the perceptually important parts of the pitch change has been captured in a closed copy stylization and it was the first step in understanding the intonation,

(Refer Slide Time: 27:23)



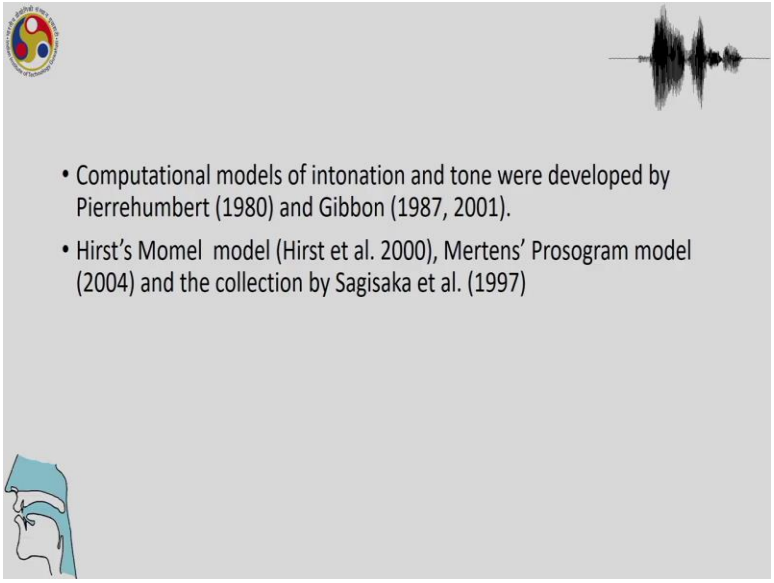
So, such a model gives us an inventory of standard pitch movements, combinations of these generate perceptually equivalent pitch contours. So, it can give us standard pitch moments that these are the pitch movements we have to look, that we have to find in an utterance, that these are the ones that will be can be predicted to be finding an utterance and also that the combinations will give you the contours.

(Refer Slide Time: 27:59)



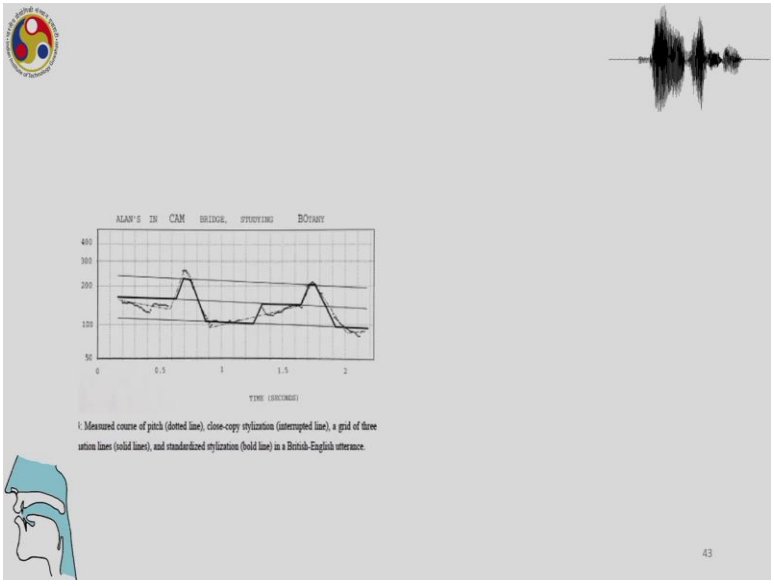
So, many approaches to competition modelling or stylization of fundamental frequency contours have been proposed. So the Fujisaki model of phrase and accent components of intonation, the linear component model of t hart et al 1990 and the Tilt model of Taylor, all these are models of competition modelling or stylization of F0.

(Refer Slide Time: 28:21)



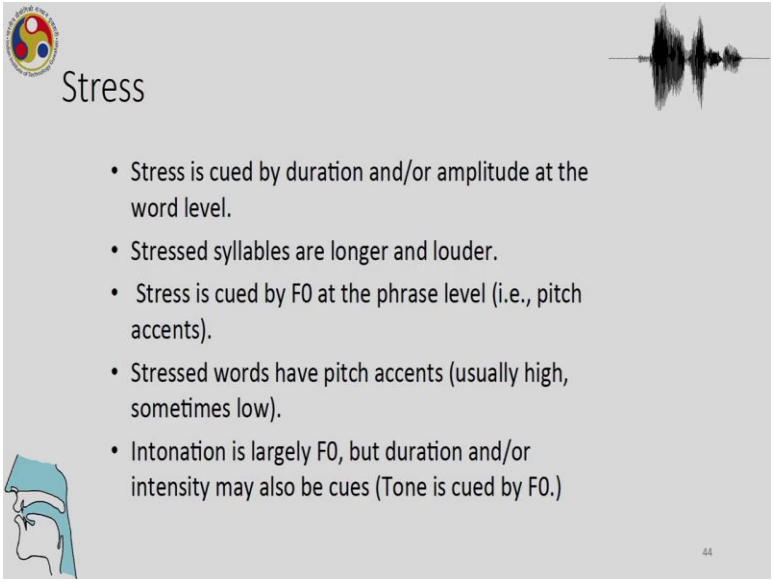
So, the competition models of intonation and tone were developed by Pierrehumbert and Gibbon, and Hirst's Momel model Mertens Prosogram model and the collection by Sagisaka et al. So, all these are computational models of tone and intonation now which have been developed.

(Refer Slide Time: 28:41)



And this is the grid lines showing the same sentence that you saw before, showing the baseline, midline and the top line that we talked about. So, we can see the baseline, you can see the slide movement for the midline and then we can see top line.

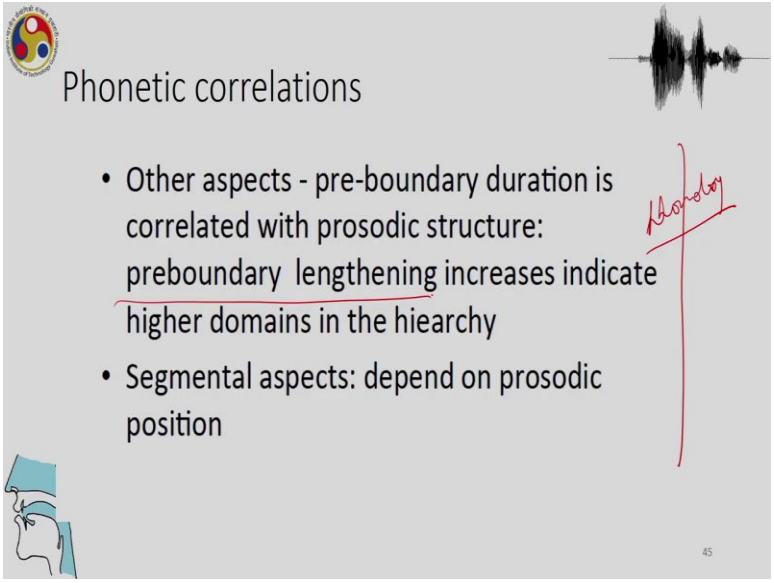
(Refer Slide Time: 29:01)



So, now, we again talk about something which we have talked about quite a lot in this course, we had an entire lecture on stress. So, stress is very important when talking about prosody and the various levels of intonation. So, stress is cued by duration and or amplitude at the word level, stress syllables are longer and louder.

Stress is cued by fundamental frequency at the phrase level at a speech accent. Remember the pitch accents, also at the word level there are the syllables in various languages. Syllables have stressed this at a higher level, at the phrase level you have sentence level. Stress words have pitch accents, intonation is largely a fundamental frequency but duration and/or intensity can also be cues.

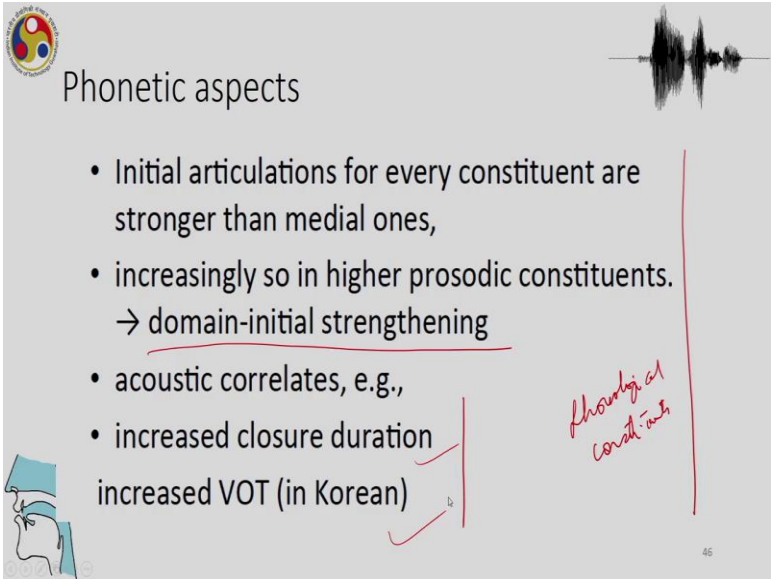
(Refer Slide Time: 30:00)



So, other aspects like pre-boundary duration is correlated with prosodic structure, pre-boundary lengthening increases indicate higher domains in the hierarchy. So, there is always boundary lengthening, showing that now when we talk about the phonology and we talk about the levels of organisation, we can understand some things about lengthening, pre-boundary lengthening et cetera, which happen at higher levels of the hierarchy where you saw.

In the beginning you saw the phonological analysis of prosody, you saw that there could be an utterance level, there could be an intonation phrase level and there could be other phonology phrase levels. And if you have pre-boundary lengthening, that is, you have two phrases, then before that you might have lengthening. Segmental aspects also depend on prosodic position sometimes.

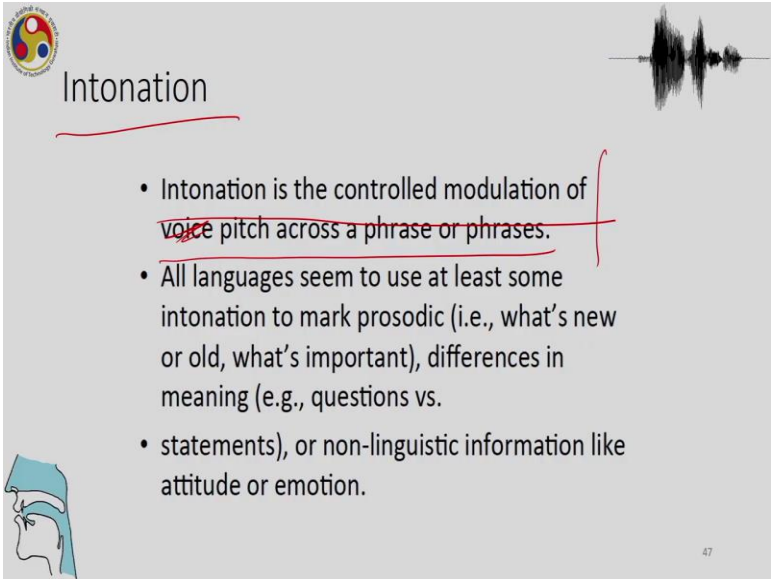
(Refer Slide Time: 31:06)



So, when we talk about segmental aspects of prosody, the one which is important which comes to mind is domain initial strengthening. And in higher prosodic consonants, again, consonants we again find domain initial strengthening, we have that is in the initial part of utterance of a phrase you will find that they are mostly financial processes like assimilation do not happen and they are segments are most faithful to their actual underlying forms and then we have strengthening in those forms as well.

And we have acoustic correlates, example, increased closure duration and increased voice onset time in higher prosodic constituents. So, these are the phonetic aspects, so these are of phonological constituents. So, we saw that how the perceptual models were developed for understanding intonation and then we have also phonological models which understand them as abstract units. When we have those abstract, even in the construction was abstract units we talk about the acoustic, phonetic aspects of prosody.

(Refer Slide Time: 32:37)

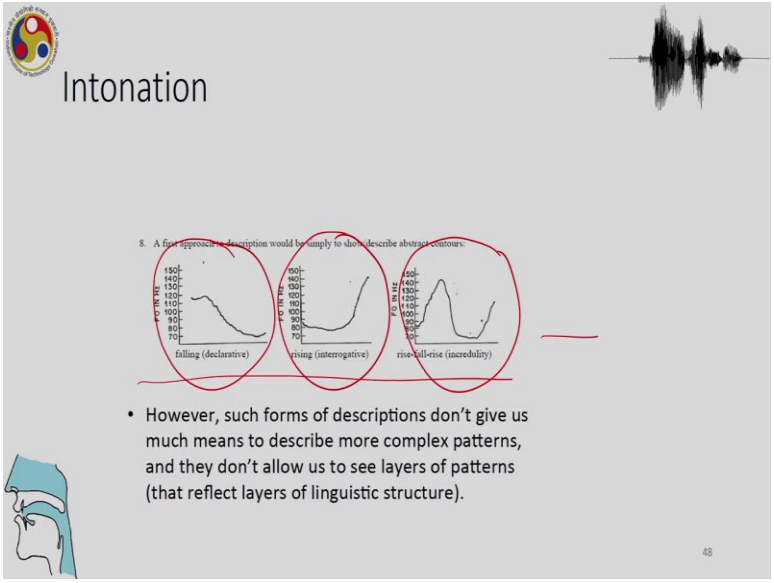


So intonation, we come to the level of the sentence level, that is the utterance when you look at it, we look at it from various angles, we look at it from the phonetic angle and also from phonological angle when you look at it. We try to understand the utterance from, we try to understand the changes. Even when we are discussing the abstract parts we try to understand how the phonetics contribute to understanding those units from the acoustic, how the acoustics contribute to our understanding of the abstract units. So, intonation is the control modulation of voice pitch across a phrase or phrases.

So, this is what, initially remember when we talked about intonation, initially we talked about how in certain interpretations we have the change in melody, that is the continuous change in relative melody was the way that is that it was defined. Now, we understood after looking at those models that actually, there is a lot of control there, when the changes happened the changes do not happen in random way. We saw how Alan's in Cambridge studying botany, so, we saw how speakers control the parts which will receive prominence.

So, that is why we call intonation a control modulation of pitch across phrase or phrases. So, all languages seem to use at least some intonation to mark prosodic differences in meaning, example questions and statements. We will see those in another lecture when we are talking about what is the difference between question and a statement or non-linguistic information like attitude or emotion.

(Refer Slide Time: 34:44)



And these are three diagrams showing you what we talked about just now that it is control modulation of pitch. So, why do we know that it is control modulation of pitch? You saw how parts were emphasised when you saw the English sentence where we saw that Alan's in Cambridge and we saw where the first part of Cambridge was emphasised. And then apart from that, what we just now said that it is controlled modulation of pitch.

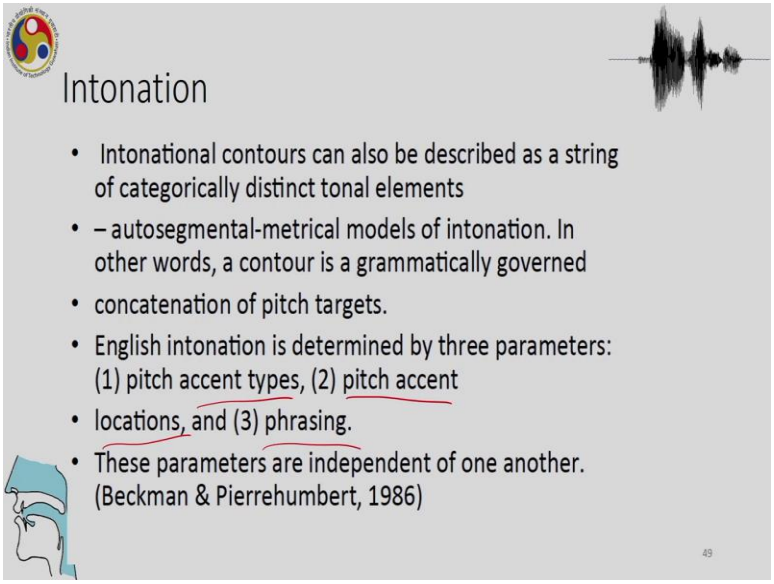
So, what do we mean by that? So, we see these three English sentences, the control modulation is seen because this is a declarative sentence, this is the pitch contour of an interrogative sentence and this is an incredulity like can you believe it. So, when you have a sentence like that, you have rise and fall and rise.

And when he was asking a question, then you have an interrogative sentence and you have a rising intonation. So to mean that you are asking a question whereas it was a falling then you it is just a declarative sentence. So, this is the control modulation of pitch and we will look at more into this in the following lecture.

And, however, once you remember that such forms of descriptions do not give us enough means to describe more complex patterns they do not allow us to see the layers of patterns. So, while this is extremely true, this is the control modulation of pitch, this does not now let us see more into this utterance.

So, we want to know how exactly, where exactly the fall happens, we want to know where exactly the rising happens, we want to know which parts are there, what are these rises and falls, where in the utterance or what is the utterance ties them to particular locations or can it be just anywhere in the sentence, can you have a rise or fall or to mean interrogative to mean falling to mean incredulity, are these tied to some locations in the sentence. So, all those complexities are not visible here, even though you can see the control modulation.

(Refer Slide Time: 37:38)



So, when we talk about intonation in the other lectures, you will see that international contours can be described as a string of categorically distinct tonal elements. And the auto-segmental metrical model that we will discuss in the other lecture shows that it is a contour, it's a grammatically governed concatenation of pitch target. And English accent an intonation is determined by three parameters, pitch, accent, types, pitch, accent and phrasing. And these parameters are independent of one another.

So, in this lecture, we have now seen how phonetics plays a role, acoustics plays a role, how our perception has been model enclosed copy stylization where we understood that some parts of an utterance are the parts where you have the actual pitch movement. And we saw that with English sentence, we saw that with the Dutch sentences. And it is important for any analysis to capture those vital points in the utterance.

But we also have models, and apart from that we have competition models which can do those predictions. And apart from these, we have phonological models which studies the complexity of an utterance in great detail as to why you have domain initial strengthening, why you have final lengthening, why you have a pause in certain parts, why do segments change at the phrase level and what exactly is governing those changes, and why do we have utterance breaks, why do we have utterances in two different chunks in a sentence that we do not when you are saying a sentence, we are actually taking breaks in the parts and why are we making those breaks.

So those details are studied in the phonological models such as auto segmental, metrical model, et cetera. And also the tones, the tones which could be called a pitch accent. And why do they occur at particular locations, what are the types, what are the brakes, so these are studied in great detail in the phonological models of intonation. And we will start looking at those in the next lecture.

So, thank you for your attention and we will continue with tone and intonation in the next lecture, where we will also talk about how a tone language is different from an intonation language. Thank you for your attention.